On the Optimal Skill Distribution in a Mirrleesian Economy

Leung, Tin Cheuk and Yazici, Hakki

Chinese University of Hong Kong, Sabanci University

June 2011

On Optimal Skill Distribution in a Mirrleesian Economy*

Tin Cheuk Leung[†]
Chinese University of Hong Kong
tleung@cuhk.edu.hk

Hakki Yazici [‡]
Sabanci University
hakkiyazici@sabanciuniv.edu

June 2, 2011

Abstract

People are heterogenous in the skills by which they turn effort into output. A central question in normative public economics is how to redistribute resources from more- to less-skilled individuals efficiently. In addition to income taxation, this paper considers another policy tool of redistribution by allowing planner to choose the dispersion of skill distribution given the average skill level of the economy. We find that, depending on the parameters of the model, either perfectly unequal skill distribution in which one group has a very high skill level and the rest are completely unskilled, or perfectly equal skill distribution in which all agents have the same skill level, is socially optimal, but an interior level of skill inequality is never optimal. We then provide conditions on the parameters under which perfectly equal and perfectly unequal skill distributions are optimal.

Keywords: Skill Distribution, Mirrleesian Taxation, Redistribution, Efficiency

JEL Classification: H2

*We would like to thank James Mirrlees for his detailed comments and encouragement. We also thank Abdurrahman Aydemir, Eren Inci, Larry Jones, Nicola Pavoni, Chris Phelan, and seminar participants at Ankara University SBF, ODTU, PET 10, TOBB-ETU, and the University of Barcelona.

[†]Contact: Department of Economics, Room 914, Esther Lee Building, Chung Chi Campus, Shatin, Hong Kong

[‡]Contact: Faculty of Arts and Social Sciences, Office 2093, Orhanli, Tuzla, 34956 Istanbul, Turkey

1 Introduction

People are heterogenous in the skills with which they turn effort into output. A central question in normative public economics is how to redistribute resources from more- to less-skilled individuals efficiently. One policy tool with which to achieve this is income taxation, redistributing resources from high- to low-income individuals. As is well-known, however, income taxation is distortionary when individuals' skills and efforts are private information (See Mirrlees (1971)).

In this paper, we consider an additional policy tool of redistribution. We do so by allowing planner to choose the dispersion of skill distribution of the economy, taking the average skill level as given.[1] By choosing a less dispersed distribution, the planner can create an economy with more equal earnings capacity among agents. This implies a more equal distribution of consumption for given income taxes. It is important, though, to realize that changing skill distribution not only affects how total output is shared across agents, but also affects the overall productivity of the economy. The amount of output that can be produced by a given labor force depends on the skill distribution chosen. Taking this effect on productivity into account, we ask how the planner should use these two policy tools jointly for the efficient redistribution of resources.

To answer this question we consider a static Mirrleesian economy in which the planner chooses the skill distribution and income taxes. In the model, the planner first chooses the skill distribution, agents then draw their types from the skill distribution privately, the planner chooses the income tax system, and finally agents work, pay taxes and consume. The main difference between our model and standard models in the optimal tax literature is the initial stage of skill distribution choice in which the planner, taking the average level of skills as given, chooses the dispersion of the skill distribution. We restrict the set of skill distributions available to the planner to discrete distributions with a finite number of mass points. The planner thus essentially chooses the value of the mass points.[2]

[1] We do not take a stance on any particular interpretation of skill distribution choice. However, one may interpret this as a choice of education policy given that people attain a significant portion of their skills through learning.

[2] We make two assumptions here. First, the number of mass points is fixed. Second, the planner takes the probability attached to each mass point as given when he chooses the value of the mass points. The latter assumption is not important because our main result holds regardless of the probability assigned to each mass point.

In such a world, at one extreme, the planner can choose a skill distribution in which the value of all mass points is equal to the average skill level. In this extreme, after the skill draw, all agents have the same earnings capacity. We call this the *perfectly equal* skill distribution. In this case, redistribution is carried out solely via skill distribution choice; there is no need for income taxation. At another extreme lies a skill distribution in which the value of all but one mass point is set to zero. In this extreme, after the draw, a fraction of agents have very high earnings capacity while the rest are completely unproductive. We call this the *perfectly unequal* skill distribution. Here, income taxes are heavily needed for redistribution. In between, there is a continuum of skill distributions available to the planner, each with a different level of skill inequality. The main result of our paper is striking: depending on the parameters of the model, either the perfectly equal or the perfectly unequal skill distribution is socially optimal, but an interior level of skill inequality is never optimal. In other words, it might be optimal to use only income taxation or egalitarian skill distribution for redistribution, but it is never optimal to use these two redistribution tools together.

In the main body of the paper, we assume that the planner faces a *linear skill constraint* with two mass points.[3] More precisely, the planner chooses mass points w_1, w_2 subject to the following skill constraint,

$$p_1 w_1 + p_2 w_2 = \alpha,$$

where p_1, p_2 are exogenous probabilities attached to the mass points and α is the average skill level in society. Under this assumption of linearity, we show that the socially optimal skill distribution is always perfectly unequal, i.e., $w_i = 0$, for some i. The intuition for this result is as follows. Suppose that $w_i > 0$, for all i. In this case, it is obvious that the optimal labor levels are positive for both types. Then, by moving to a skill distribution in which the type with a higher labor level has all the skills and setting the labor level of the other type to zero, the planner increases total output and decreases total disutility. This shows that increasing skill inequality benefits society because it increases productive efficiency. Under full information, income taxes are not distortionary, which means that the planner can distribute consumption according to its will using income taxes at no cost. This implies that the productive efficiency

[3]The assumption of two mass points is not restrictive. We show that all of our results hold in the case with N types.

gain is the only effect of increasing skill inequality on the economy. Thus, under full information, perfectly unequal skill distribution is socially optimal. However, when skill levels are private information, income taxes are distortionary. Increasing skill inequality exacerbates the distortion associated with income taxation, because the benefit of pretending to be low-skilled is higher for the high-skilled agents when skill inequality is higher. Therefore, increasing skill inequality imposes a cost on society as well. We show, however, that the socially optimal skill distribution is still perfectly unequal.

Next, we question the robustness of the optimality of perfectly unequal skill distribution. To do so, we analyze several extensions of the baseline model, and find that the optimal skill distribution is either perfectly unequal or perfectly equal, depending on the parameters of the model. In the first extension, we relax the linear skill constraint assumption by allowing the skill constraint to be convex in skill levels as follows,

$$p_1 w_1^{\beta} + p_2 w_2^{\beta} = \alpha,$$

where $\beta \geq 1$ is the convexity parameter. We show that under full information, the socially optimal skill distribution is either perfectly unequal or perfectly equal, depending on the convexity of the skill constraint and the convexity of the disutility function. When there is private information about skills, we provide a sufficient condition for the optimality of perfectly equal skill distribution. When this condition does not hold, it is hard to provide an analytical solution. Instead, we parameterize the utility and disutility functions, and solve the planner's problem numerically. We show that the socially optimal skill distribution is again either perfectly equal or perfectly unequal. This time, the concavity of the utility function is also one of the solution's determinants.

Second, we extend the baseline model by putting an exogenous uniform lower bound on the skill levels of agents. The lower bound can be interpreted as skills that agents are born with. In this case, we compute the socially optimal allocation numerically and find very robustly that the optimal skill distribution is again extreme. Furthermore, we find that if the lower bound is low relative to the average skill level, then the optimal skill distribution is the perfectly unequal one. Finally, we extend the set of skill distributions available to the planner to discrete distributions with three mass points to show that our results do not depend on the number of types allowed in the skill distribution. We show, as in the baseline model, that the optimal skill distribution is again perfectly unequal, whereby the values of two mass points are set to zero and the value of the remaining mass point to the maximum.

The paper closest to ours is Cremer, Pestieau, and Racionero (2010), which has a similar model setup to our baseline model. However, their paper only compares the two extreme skill distributions, perfectly equal and perfectly unequal, and shows that the perfectly unequal distribution provides higher social welfare under a linear skill constraint. So, their paper only shows that perfectly equal skill distribution cannot be the solution to the social planning problem. Our paper considers the whole set of feasible skill distributions and proves that perfectly unequal skill distribution provides the highest social welfare, and hence, is the socially optimal distribution. Furthermore, in the case of convex skill constraints, our paper provides conditions on the parameters under which perfectly equal and perfectly unequal skill distributions are socially optimal whereas Cremer, Pestieau, and Racionero (2010) again merely compares perfectly equal and unequal skill distributions.

A few other papers analyze comparative static properties of optimal allocations with respect to certain parameters of the skill distribution. Instead of providing a comprehensive survey of this literature, we will provide two brief examples.[4] Brett and Weymark (2008) investigates the effect of changing an agent's skill level on the solution of a Mirrlees optimal income tax problem. Hamilton and Pestieau (2005) studies the effect on individual utilities of changing the fraction of individuals when the social welfare function is either maximin or maximax. The main difference between these papers and ours should be apparent: while we analyze the *optimal* skill distribution, they only study comparative static properties.

In our analysis we do not take a stance on any particular interpretation of the skill distribution choice. However, if we think that skills can be partly attained through education, our model may have implications for education policy. In this regard, the paper is related to several papers that consider education policy as a redistribution tool in the presence of income taxation. Hare and Ulph (1979) shows that when agents' learning abilities are heterogenous and skill types

[4]Other papers that provide such comparative static results are Boadway and Pestieau (2006), and Simula (2007).

are observable, the "optimal choice of education policy reinforces [the] redistributive effect of income tax." Bovenberg and Jacobs (2006) constructs a model in which agents choose the education level. The government can only influence education, and thus skill distribution, through education subsidies. They show that providing more subsidies to smarter agents (a regressive education policy) would make it more incentive compatible to undertake more redistribution through income taxes.

Another potential interpretation of skill distribution choice is related to Skill-Biased Technical Change (SBTC), which refers to "a shift in the production technology that favors skilled over unskilled labor by increasing its relative productivity" (Violante (2009)). Under this interpretation, the planner chooses the level of SBTC that determines the level of skill inequality between skilled and unskilled labor. For instance, suppose that the government is choosing the degree of computerization of the production process. In this case, a higher degree of computerization increases the productivity of some workers (those who are more prone to using computers) but decreases the productivity of the rest, thereby changing the skill distribution towards higher inequality. While most papers in the SBTC literature are positive studies of the growth and income distribution implications of SBTC, this paper could be interpreted as a normative analysis of the optimal level of SBTC. However, this interpretation of our model should be approached with caution because we treat skilled and unskilled labor as perfect substitutes as in almost all the Mirleesian taxation literature (see Naito (1999) and Stiglitz (1982) for exceptions), contrary to the empirically relevant case of production functions with complementarity.

The rest of this paper is structured as follows. In Section 2, we introduce the model formally. Section 3 analyzes the optimal skill distribution problem. Section 4 discusses the extensions of the baseline model. Finally, Section 5 concludes the paper.

2 Model

There is a unit measure of agents. They produce output individually according to the production function

$$y = wl,$$

where y denotes output, w denotes skill level, and l denotes labor effort.

Each agent's preference is given by

$$u(c) - v(l),$$

where c is consumption and u and v satisfy $u', -u'', v' > 0$ and $v'' > 0$.

The novelty of our analysis is that we allow the planner to choose the distribution of skills. For tractability, we assume that the planner has to choose a distribution in which skills can take only two values, w_1 and w_2. The probability of drawing w_1 is p_1 and the probability of drawing w_2 is p_2. We allow the planner to choose w_1 and w_2, but p_1 and p_2 are exogenously given. We take the average skill level of the economy as given, at α. We assume that the planner chooses w_1 and w_2 subject to a linear skill constraint:

$$p_1 w_1 + p_2 w_2 \leq \alpha.$$

The constraint states that the average skill level of the distribution chosen by planner cannot exceed α.

Allocation. An *allocation* in this economy is defined as $(w_i, c_i, l_i)_{i=1,2}$, where c_i and l_i represent consumption and labor allocation of type i.

Feasibility. An allocation is *feasible* if

$$p_2 c_2 + p_1 c_1 \leq p_2 w_2 l_2 + p_1 w_1 l_1, \tag{1}$$

$$p_1 w_1 + p_2 w_2 \leq \alpha, \tag{2}$$

$$w_1, w_2, c_1, c_2, l_1, l_2 \geq 0. \tag{3}$$

The first inequality here states that total consumption cannot exceed total output. The second inequality makes sure that the average skill level of the distribution chosen by planner does not exceed α. Finally, the third inequality is just the non-negativity of skill, consumption and labor allocations.

The timing of the events is as follows. First, the planner chooses the skill distribution. Then, each agent privately draws her skill from this distribution. Finally, the planner chooses the consumption and labor allocations, agents announce their types and receive the corresponding allocation. This informational friction requires the allocation to satisfy the following familiar *incentive compatibility* conditions:

Incentive compatibility. An allocation is *incentive compatible* if

$$u(c_2) - v(l_2) \geq u(c_1) - v(w_1 l_1 / w_2) \tag{4}$$

$$u(c_1) - v(l_1) \geq u(c_2) - v(w_2 l_2 / w_1) \tag{5}$$

A social planner chooses the level of consumption, labor and the skill distribution to maximize total welfare subject to social feasibility and incentive compatibility constraints.

Social Optimum. An allocation is a *social optimum* if it solves[5]

$$\max_{w_1, w_2, c_1, l_1, c_2, l_2} p_2[u(c_2) - v(l_2)] + p_1[u(c_1) - v(l_1)]$$

s.t. (1), (2), (3), (4), and (5).

We denote the optimal allocation by $(w_1^*, w_2^*, c_1^*, l_1^*, c_2^*, l_2^*)$.

As we are interested in the socially optimal skill distribution, we focus on w_1^* and w_2^* in the above problem. To understand the question at hand, it is helpful to consider the set of distributions that are available to society. On the one extreme, we can set $w_1 = 0$ and $w_2 = \frac{\alpha}{p_2}$, or $w_1 = \frac{\alpha}{p_1}$ and $w_2 = 0$. In both of these cases, a fraction of agents have very high earnings capacity while the rest are completely unproductive. We call these perfectly unequal skill distributions. On the other extreme, we can set $w_1 = w_2 = \alpha$ and make everyone in the society identical. We call this the perfectly equal skill distribution. In between, there is a whole range of skill distributions in which both $w_1, w_2 > 0$. In some of these distributions, $w_1 > w_2$ and in some $w_1 < w_2$.

From now on, we denote by H the type that the planner allocates higher skills and by L the other type, i.e., $w_i = w_H$ and $w_j = w_L$, if $w_i > w_j$. In addition, let $p_i = p_H$ and $p_j = p_L$.

[5] We use a utilitarian social welfare function with equal weights on all agents. However, all of our results hold under any social welfare function that values equality beyond the laissez-faire market outcome. The only feature of this utilitarian social welfare function on which we rely is that the high-skilled type's incentive constraint binds, which is true under any social welfare function that values equality.

Hence, we redefine an allocation as $(w_H, w_L, c_H, l_H, c_L, l_L)$. In section 3, we show that when skill constraint is linear, the socially optimal skill distribution involves perfect inequality, meaning $w_L = 0$ and $w_H = \frac{\alpha}{p_H}$. Then, in section 4, we extend the model in several dimensions and prove that either perfect inequality or perfect equality is optimal depending on certain critical parameter values.

2.1 Rewriting Planner's Problem

Let $\theta = \frac{w_L}{w_H}$. Observe that $\theta = 0$ is the case in which there is perfect inequality in skill distribution. As we increase θ towards 1, inequality in skill distribution decreases and at $\theta = 1$ there is perfect equality of skills. In the rest of the paper, we will be interested in the value of socially optimal θ.

It is a well-known result that only the type H incentive constraint binds under a Utilitarian social welfare function with equal weights. Now we can substitute the skill constraint into the resource constraint and rewrite the problem as:

$$\max_{\theta, c_L, l_L, c_H, l_H} p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]$$

s.t.

$$p_H c_H + p_L c_L \leq \alpha l_H + \frac{\alpha p_L \theta (l_L - l_H)}{p_H + p_L \theta}$$

$$u(c_H) - v(l_H) \geq u(c_L) - v(\theta l_L)$$

$$c_L, c_H, l_L, l_H \geq 0$$

$$\theta \in [0, 1].$$

If the planner sets $\theta = 1$, then agents choose their types from the perfectly equal skill distribution where all agents have the skill level α. In this case, the right-hand side of feasibility becomes $p_H \alpha l_H + p_L \alpha l_L$ and the incentive compatibility constraint disappears.

3 Socially Optimal Skill Distribution

3.1 Benchmark: Full Information Social Optimum

We first analyze the benchmark case with full information. The planner's problem is the same as above except that there is no incentive compatibility constraint.

It is always optimal for the planner in the full information case to choose the perfectly unequal skill distribution, i.e. giving all of the skill to one of the two types. The intuition is straightforward. In the absence of the incentive constraint, the planner can always equate consumption across agents at zero cost, so the only criterion of optimal skill distribution is productive efficiency. It is easy to see that productive efficiency requires skill distribution to be perfectly unequal. Suppose that this is not true, $\theta^* \in (0, 1]$. From the first order optimality condition between labor of H and L, we have:

$$v'(l_L^*) = \theta^* v'(l_H^*)$$

which implies that $l_H^* \geq l_L^*$ since $v'' > 0$ and $\theta^* \in (0, 1]$. We can find another feasible allocation that strictly improves welfare. Consider a new allocation in which $\hat{\theta} = 0$, $\hat{l_L} = 0$, and the rest of the allocation stays the same. The feasibility constraint is relaxed because

$$\alpha l_H^* + \frac{\alpha p_L \hat{\theta}(\hat{l_L} - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^* \geq \alpha l_H^* + \frac{\alpha p_L \theta^*(l_L^* - l_H^*)}{p_H + p_L \theta^*}$$

and the disutility of L type decreases because

$$v(\hat{l_L}) = v(0) < v(l_L^*)$$

which indicates that any $\theta \in (0, 1]$ cannot be optimal. Therefore, we have the following theorem:

Theorem 1. *In the full information social optimum with linear skill constraint, $\theta^* = 0$.*

3.2 Private Information Social Optimum

With private information, the choice of skill distribution not only affects productive efficiency, but also, through the incentive constraint, affects the set of consumption distributions available to the planner. In this section, we show that if the skill constraint is linear, then the optimal skill distribution is still perfectly unequal.

First consider the effect of skill distribution choice on the production side of the economy. From the analysis of the full information case, we know that increasing skill inequality increases total output and decreases total disutility at the same time. Therefore, productive efficiency pushes towards the perfectly unequal skill distribution. However, when skill is private information, increasing skill inequality increases the distortions associated with income redistribution. To see this, consider the incentive constraint of type H which holds with equality at the optimal allocation:

$$u(c_H) - v(l_H) = u(c_L) - v(\theta l_L).$$

When the planner increases skill inequality, meaning a decrease in θ, keeping the rest of the allocation intact would violate the incentive constraint. This means that the planner has to accompany the increase in skill inequality by increasing consumption inequality and/or by increasing l_L relative to l_H. Both are distortionary and involve a cost to society. Therefore, unlike the full information benchmark, increasing skill inequality not only has a productive efficiency gain but it also has a cost in terms of increasing the distortions associated with income redistribution. Theorem 2 formally proves that the optimal skill distribution is still perfectly unequal.

Theorem 2. *In the private information social optimum, $\theta^* = 0$.*

Proof. We proceed in two steps.

Step 1: θ^* cannot be interior.

Suppose not, $\theta^* \in (0, 1)$, and there are two cases to consider.

Case 1: $l_H^* \geq l_L^*$

Consider a new allocation where $\hat{\theta} = 0$, $\hat{l_H} = l_H^*$, $\hat{l_L} = 0$, $\hat{c_H} > c_H^*$ and $\hat{c_L}$ such that the incentive constraint still holds,

$$u(\hat{c_H}) - v(l_H^*) = u(\hat{c_L}) - v(0)$$

,

and total output is used up for consumption.

To see that we can make the above equality hold in the new allocation, observe the following. In the new allocation, total output is weakly higher because it is given by

$$\alpha[l_H^* + \underbrace{\frac{p_L\theta(l_L^* - l_H^*)}{p_H + p_L\theta}}_{\leq 0}]$$

which is decreasing in θ. Now if we gave all of the output to the high type, he would get $\frac{\alpha}{p_H} l_H^* > c_H^*$.

We can show that $u(\frac{\alpha}{p_H} l_H^*) - v(l_H^*) > u(0) - v(0)$ because if not, we have

$$u(0) - v(0) > u(\frac{\alpha}{p_H} l_H) - v(l_H^*) \quad > \quad u(c_H^*) - v(l_H^*),$$

where the second inequality is true because $\frac{\alpha}{p_H} l_H^* > c_H^*$. Also observe that $u(0) - v(0) > u(c_H^*) - v(l_H^*) \geq u(c_L^*) - v(\theta l_L^*) > u(c_L^*) - v(l_L^*)$. Therefore, the $c_H = c_L = l_H = l_L = 0$ allocation gives a strictly higher utility than the efficient allocation, which is clearly incentive compatible and feasible, so efficient allocation cannot be welfare maximizing, which is a contradiction.

Hence, we have shown that $u(\frac{\alpha}{p_H} l_H^*) - v(l_H^*) > u(0) - v(0)$. Now, decrease c_H and increase c_L until this inequality holds with equality, and that is how we construct $\hat{c_H}$ and $\hat{c_L}$.

Finally, observe that $\hat{c_H} > c_H^*$, because otherwise $\hat{c_L} > c_L^*$, and that would mean

$$u(\hat{c_H}) - v(l_H^*) \leq u(c_H^*) - v(l_H^*) = u(c_L^*) - v(\theta l_L^*) < u(\hat{c_L}) - v(0),$$

which is a contradiction.

$\hat{c_H} > c_H^*$ implies that H type's welfare is strictly higher in the new allocation and the incentive constraint holding with equality implies that so is L type's. Hence, the new allocation strictly improves over the efficient allocation, yielding the desired contradiction.

Case 2: $l_H^* < l_L^*$

We can set $\hat{\theta} = 1$, then the resource constraint is relaxed because

$$\alpha l_H^* + \frac{\alpha p_L \hat{\theta}(l_L^* - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^* + \frac{\alpha p_L (l_L^* - l_H^*)}{p_H + p_L} > \alpha l_H^* + \frac{\alpha p_L \theta^* (l_L^* - l_H^*)}{p_H + p_L \theta^*}$$

The incentive constraint is also relaxed because

$$u(c_H^*) - v(l_H^*) = u(c_L^*) - v(\theta^* l_L^*) > u(c_L^*) - v(\hat{\theta} l_L^*) = u(c_L^*) - v(l_L^*)$$

Thus, it is easy to find another allocation that improves welfare.

Step 2: $\theta^* = 0$

Suppose not, $\theta^* = 1$, then it is easy to show $c_H^* = c_L^*$ and $l_H^* = l_L^*$, but then we can set $\hat{\theta} = 0$ and $\hat{l_L} = 0$. The resource constraint is unaffected because

$$\alpha l_H^* + \frac{\alpha p_L \theta^* (l_L^* - l_H^*)}{p_H + p_L \theta^*} = \alpha l_H^* + \frac{\alpha p_L \hat{\theta}(\hat{l_L} - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^*$$

We can also set $\hat{c_H} > c_H^*$ and $\hat{c_L}$ such that the incentive constraint holds, as we did in Step 1 Case 1:

$$u(\hat{c_H}) - v(l_H^*) = u(\hat{c_L}) - v(0)$$

This improves welfare because $\hat{c_H} > c_H^*$ and $\hat{l_L} = 0 < l_L^*$. □

4 Extensions

In the foregoing analysis, we show that the optimal skill distribution is extreme, in particular $\theta^* = 0$. In this section, we generalize the baseline model in several dimensions to check how robust is the optimality of perfectly unequal skill distribution. We find that the optimal skill distribution is either perfectly unequal or perfectly equal, depending on the parameters of the model.

4.1 Convex Skill Constraint

In the previous analysis, we analyze a special case in which the skill constraint is linear. In reality, however, there can be decreasing returns to scale when we transfer skills from the low type to the high type. To allow for this possibility, we now assume a convex skill constraint as follows:

$$p_L w_L^{\beta} + p_H w_H^{\beta} \leq \alpha \tag{6}$$

where $\beta \geq 1$ is the scalar for convexity of the skill constraint.

Using this skill constraint, we can write the total output as:

$$\begin{aligned} p_L w_L l_L + p_H w_H l_H &= w_H(\theta p_L l_L + p_H l_H) \\ &= (\frac{\alpha}{p_L\theta^{\beta} + p_H})^{1/\beta}(\theta p_L l_L + p_H l_H) \end{aligned}$$

4.1.1 Full Information

First, we analyze the benchmark case with full information. Using the above expression for total output, we can write the planner problem as:

$$\max_{\theta, c_H, c_L, l_H, l_L} p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]$$

s.t.

$$p_H c_H + p_L c_L = (\tfrac{\alpha}{p_L\theta^{\beta}+p_H})^{1/\beta}(\theta p_L l_L + p_H l_H)$$
$$c_H, c_L, l_H, l_L \geq 0, \theta \in [0, 1].$$

Denote $c_H^*(\theta), c_L^*(\theta), l_H^*(\theta), l_L^*(\theta)$ as the values of c_H, c_L, l_H, l_L that maximize the above problem for a given θ, and denote U^* as the maximized total utility:

$$U^* \equiv p_H[u(c_H^*(\theta)) - v(l_H^*(\theta))] + p_L[u(c_L^*(\theta)) - v(l_L^*(\theta))]$$

We are interested in the optimal value of θ that maximizes U^*. The Envelop Theorem implies that the total derivative of the total utility with respect to θ can be expressed as the product of the multiplier to the resource constraint (λ) and the partial derivative of total output with respect to θ:

$$\begin{aligned}\frac{dU^*}{d\theta} &= \frac{\partial\lambda[(\frac{\alpha}{p_L\theta^\beta+p_H})^{1/\beta}(\theta p_L l_L^*(\theta)+p_H l_H^*(\theta))-p_H c_H^*(\theta)-p_L c_L^*(\theta)]}{\partial\theta} \\ &= \lambda\frac{\alpha^{1/\beta}p_H p_L}{(\theta p_L l_L^*(\theta)+p_H l_H^*(\theta))^{1/\beta+1}}[l_L^*(\theta)-\theta^{\beta-1}l_H^*(\theta)]\end{aligned}$$

This expression states that whether more skill equality (higher θ) can increase welfare depends on whether it can increase total output. Productive efficiency is the only concern because, under full information, no incentive constraint restricts the planner from equalizing consumption. Using the above expression, whether total output is increasing or decreasing in θ depends on the sign of the following expression:

$$[l_L^*(\theta)-\theta^{\beta-1}l_H^*(\theta)] = l_L^*(\theta)[1-\frac{\theta^{\beta-1}}{l_L^*(\theta)/l_H^*(\theta)}]$$

This expression tells us that whether output increases as skill inequality increases (θ decreases) depends on two parameters: β and the convexity of the disutility function. The intuition is as follows. First, holding the $\frac{l_L^*}{l_H^*}$ ratio constant, when $\beta = 1$, increasing skill inequality means transferring skills to the type that is already working harder ($l_L^* < l_H^*$), keeping the average skill level constant. This increases total output. However, when $\beta > 1$, increasing skill inequality is costly because it decreases the average skill level in the economy. In fact, this cost becomes larger as β increases, and hence increasing skill inequality is less likely to increase total output for higher β. Second, the convexity of the disutility is important because it determines the $\frac{l_L^*}{l_H^*}$ ratio. If disutility is more convex, then $\frac{l_L^*}{l_H^*}$ is closer to one, and hence increasing skill inequality is less likely to increase total output.

To simplify the analysis, we assume a particular form for the disutility function, $v(l) = l^\gamma$. Then, the above expression becomes

$$[l_L^*(\theta)-\theta^{\beta-1}l_H^*(\theta)] = l_L^*(\theta)[1-\theta^{\beta-1+\frac{1}{1-\gamma}}]$$

and we have the following result.

Theorem 3. *In the full information social optimum,*

$$\theta^* = \begin{cases} 0 & \textit{if } \beta - 1 + \frac{1}{1-\gamma} < 0; \\ \textit{[0,1]} & \textit{if } \beta - 1 + \frac{1}{1-\gamma} = 0; \\ 1 & \textit{if } \beta - 1 + \frac{1}{1-\gamma} > 0. \end{cases}$$

The theorem states that optimal skill distribution is either perfectly equal or unequal, depending on the values of β and γ. In particular, as our intuition suggested in the case with the general disutility function, as β and γ increase it is more likely for the perfectly equal skill distribution to be optimal.

4.1.2 Private Information

Under private information, the planning problem is the same as the full information planning problem, except that there is an additional constraint– the usual incentive compatibility constraint.

$$u(c_H) - v(l_H) \geq u(c_L) - v(\theta l_L).$$

We know from the full information analysis of the previous subsection that productive efficiency calls for the optimality of the perfectly equal skill distribution whenever $\beta - 1 + \frac{1}{1-\gamma} \geq 0$. As the perfectly equal skill distribution also brings perfect consumption equality without any further need for distortionary income taxation, it is also socially optimal under private information whenever this condition is satisfied. The theorem below follows.

Theorem 4. *If* $\beta - 1 + \frac{1}{1-\gamma} \geq 0$*, then we have* $\theta^* = 1$ *in the private information social optimum.*

Proof. Consider a relaxed version of the private information planning problem in which we drop the incentive constraint. That relaxed problem is equivalent to the full information planning problem and we know that in the solution to that problem we have $\theta = 1$, $c_H = c_L$, and $l_H = l_L$. Clearly, this allocation satisfies the incentive constraint and hence is in the constraint set of the original planning problem under private information, which means that it solves the problem. □

It is difficult to provide an analytical solution to this problem when $\beta - 1 + \frac{1}{1-\gamma} < 0$. Therefore, in what follows we provide numerical solutions. We use constant relative risk aversion utility function

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}.$$

We compute the value of the planning problem for $\theta \in [0,1]$, and plot these value functions in Figure 1 for different parameter values. As the concavity of the utility function and convexity of the disutility function affect the result, we compute the value function with two values of σ ($\sigma = 0.5$ and $\sigma = 3$) and two values of γ ($\gamma = 2$ and $\gamma = 4$). The four sub-figures of Figure 1 each correspond to one of the four combinations of σ and γ. To illustrate the importance of the convexity of the skill constraint on the result, we compute the value function for 4 different values of β. Within each sub-figure, each plot corresponds to a value function for a different value of β.

The main message of Figure 1 is that there is no combination of parameters (β, γ, σ) for which optimal θ is interior. Several factors determine the corner in which the optimal θ will lie. The first factor is the convexity of the skill constraint. As the intuition in the full information case suggests, a more linear skill constraint (lower β) makes unequal skill distribution less costly and hence more likely to be socially optimal. Figure 1 confirms this: fixing γ and σ (i.e. looking at each sub-figure), θ^* changes from 0 to 1 as β increases. Also, observe that as we prove in section 3, when $\beta = 1$, $\theta^* = 0$, which is true in all four sub-figures.

Second, again similar to the full information case, a more convex disutility (higher γ) implies that $\frac{l_L^*}{l_H^*}$ has to be higher. Thus, increasing skill inequality (lower θ) increases output less, and hence is less likely to be optimal. In Figure 1, holding σ and β constant, a higher γ (from 2 to 4) makes $\theta^* = 1$ more likely. For instance, when $\sigma = 0.5$ and $\beta = 1.5$, θ^* changes from 0 to 1 when γ goes up from 2 to 4.

The third parameter that matters for the optimal θ is σ, the concavity parameter of the utility function. Under private information, for any σ, if the planner wants the high type to produce more output, he must provide the high type with incentives to do that, meaning $c_H > c_L$. Now, keeping everything else constant, if we look at an economy with a higher σ, the planner would like to set consumption levels of the two types closer to each other. To do this, however, the planner has to close the gap between l_H and l_L, which makes skill equality more appealing. In Figure 1, holding β and γ constant, a higher σ (from 0.5 to 3) makes $\theta^* = 1$ more likely. For instance, when $\gamma = 2$ and $\beta = 1.5$, θ^* changes from 0 to 1 when σ goes up from 0.5 to 3.

4.2 Additive innate skills

We showed in section 3 that the socially optimal skill distribution is perfectly unequal, meaning $\theta^* = 0$. However, one might think that all the people in an economy are born with a certain level of innate skills that the planner cannot grab away from them. We incorporate this into our analysis by considering the baseline model with linear skill constraint in which everyone has $K > 0$ units of innate skills. Hence the skill level of type L is $w_L + K$ and the skill level of type H is $w_H + K$, where $p_l w_L + p_H w_H = \alpha$. The question we ask in this subsection is whether the socially optimal skill distribution is still the perfectly unequal one, meaning $\theta^* = \frac{w_L^*}{w_H^*} = 0$.

Define $\tilde{\theta} = \frac{w_L + K}{w_H + K}$ and $\tilde{\alpha} = \alpha + K$. Then, we can write the feasibility in the form similar to before:

$$p_H c_H + p_L c_L \leq \tilde{\alpha} l_H + \frac{\tilde{\alpha} p_L \tilde{\theta}(l_L - l_H)}{p_H + p_L \tilde{\theta}}. \tag{7}$$

Incentive compatibility becomes:

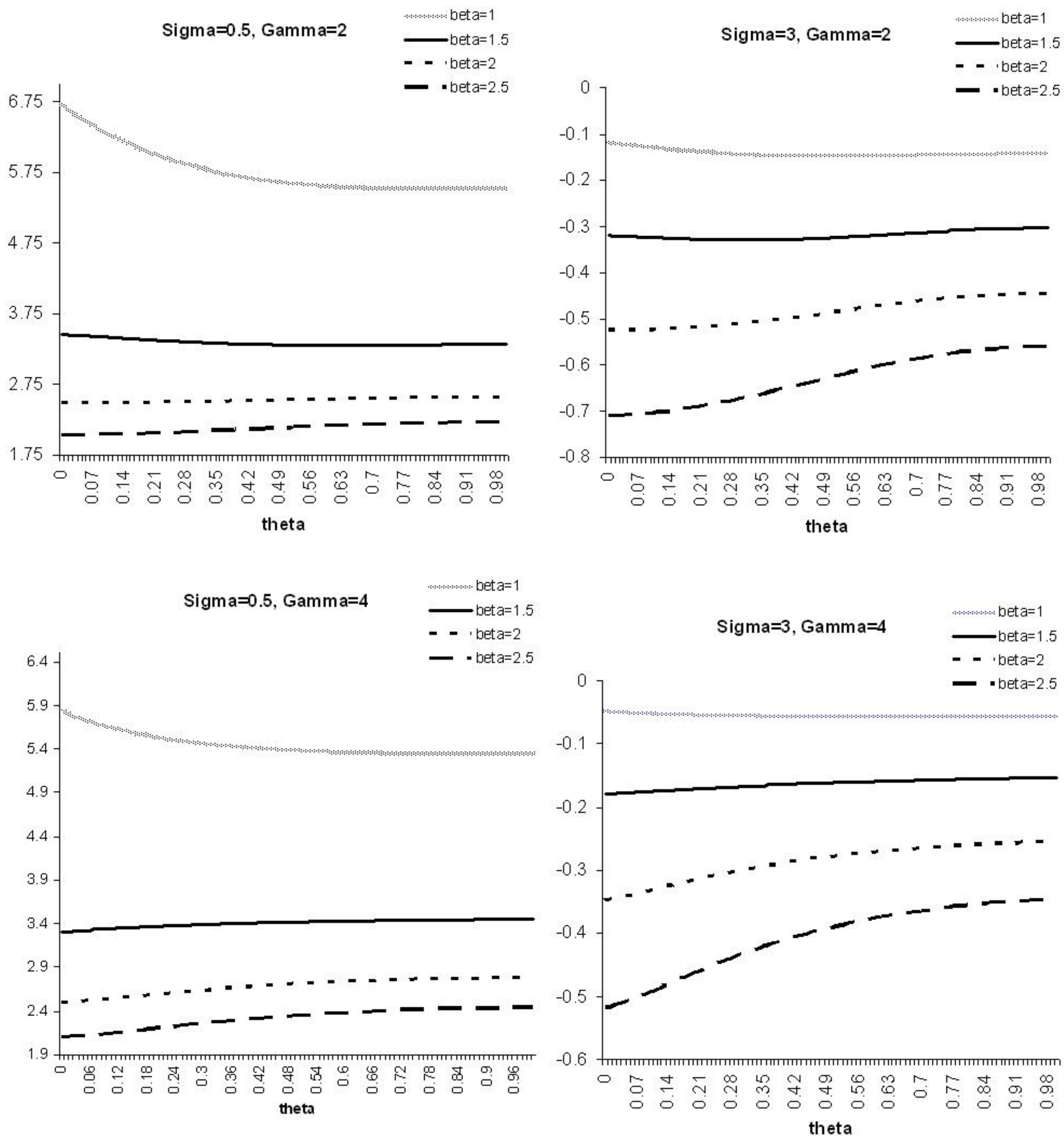
$$u(c_H) - v(l_H) \geq u(c_L) - v(\tilde{\theta} l_L).$$

This problem is exactly the same as the problem in the baseline model where the average skill in society is $\tilde{\alpha}$ and the planner is restricted to choosing among skill distributions in which each mass point has to be greater than or equal to K. This means that the domain of $\tilde{\theta}$ for the planner is

$$\tilde{\theta} \in [\frac{K}{\alpha/p_H + K}, 1].$$

In all four sub-figures of Figure 1 we can see that even in the extended model with innate skills and linear skill constraint ($\beta = 1$), an interior distribution is never optimal. In particular, when K is low relative to α, the lower bound on $\tilde{\theta}$ is low and the figures tell us that we still

Figure 1:



Value Functions

have $\theta^* = 0$ ($\tilde{\theta}^* = \frac{K}{\alpha/p_H + K}$) to be socially optimal. When K is high enough relative to α, the lower bound on $\tilde{\theta}$ is high and hence θ^* can be equal to 0 or 1, depending on the concavity of the utility function. In a similar fashion, one can also show that the extreme result of either $\theta^* = 0$ or $\theta^* = 1$ is robust to introducing innate abilities to the convex skill constraint model.

4.3 Three types

So far we have assumed that the planner is allowed to choose discrete distributions with two mass points. In this subsection, we ask if the extremity of the socially optimal skill distribution depends on this assumption. We analyze the baseline model of section 3 and allow the planner to choose discrete distributions with three mass points. As in section 3, we find that the perfectly unequal skill distribution, in which one type receives all the skills and the rest receive none, is socially optimal.

Planning Problem:

$$\max_{w_1,w_2,w_3,c_1,l_1,c_2,l_2,c_3,l_3} p_3[u(c_3) - v(l_3)] + p_2[u(c_2) - v(l_2)] + p_1[u(c_1) - v(l_1)]$$

s.t.

$$p_3c_3 + p_2c_2 + p_1c_1 \leq p_3w_3l_3 + p_2w_2l_2 + p_1w_1l_1, \tag{8}$$

$$p_1w_1 + p_2w_2 + p_3w_3 \leq \alpha, \tag{9}$$

$$w_1, w_2, w_3, c_1, c_2, c_3, l_1, l_2, l_3 \geq 0. \tag{10}$$

$$u(c_i) - v(l_i) \geq u(c_j) - v(\tfrac{w_j}{w_i}l_j), \text{for i,j=1,2,3.} \tag{11}$$

As we did previously, by H we denote the type that receives the highest skills in the distribution chosen by the planner, by M the type that receives medium skills, and by L the other type, i.e., $w_i = w_H$, $w_j = w_M$, and $w_k = w_L$, if $w_i > w_j > w_k$. In addition, $p_i = p_H$, $p_j = p_M$, and $p_k = p_L$. Hence, redefine an allocation as $(w_H, w_M, w_L, c_H, l_H, c_M, l_M, c_L, l_L)$. We also define $\theta_M = \frac{w_M}{w_H}$ and $\theta_L = \frac{w_L}{w_H}$. The following theorem provides the formal result.

Theorem 5. *Both under full information and private information, $\theta_L^* = \theta_M^* = 0$ is socially* *optimal.*

Proof. The proof for the full information case is exactly the same as the proof of Theorem 1. Hence, we provide the proof of the private information case only.

Suppose for contradiction that both θ_L^* and θ_M^* are strictly positive (other cases are straightforward implications of the two-type case).

There are three possibilities.

Case 1. $l_H^* \geq l_i^*$ for $i = M, L$.

Consider a new allocation in which we set $\theta_L = \theta_M = 0$, $l_L = l_M = 0$, and $l_H = l_H^*$. This weakly increases total output and strictly decreases total disutility. We construct the consumption part of the new allocation in the following way. Set $c_M = c_L$, and $u(c_H) - v(l_H^*) = u(c_M) - v(0)$. To show that such a consumption allocation is achievable, we first show $u(\alpha l_H^*) - v(l_H^*) > u(0) - v(0)$. Suppose not. Then, $u(c_H^*) - v(l_H^*) \leq u(0) - v(0)$, which implies $u(c_L^*) - v(l_L^*) < u(c_M^*) - v(l_M^*) < u(c_H^*) - v(l_H^*) \leq u(0) - v(0)$, which is a contradiction because then $c_i = l_i = 0$ for all i allocation dominates the efficient allocation. Hence, in the new allocation we have $u(c_H) - v(l_H^*) = u(c_M) - v(0) = u(c_L) - v(0)$. If $c_H \geq c_H^*$, then type H is better off, which further implies type L and type M are strictly better off. This means that the new allocation is an improvement over the efficient one, which is a contradiction. If $c_H < c_H^*$, then another consumption allocation in which we set c_H equals to c_H^* and divide the rest of the output equally between type L and type M provides a strictly higher total welfare compared to the efficient allocation because type H's welfare is unchanged and type L and type M's average welfare strictly increases (as the total output they consume is higher and they share it equally). This allocation is obviously incentive compatible, which again means that it is an improvement over the efficient allocation, once again a contradiction.

Case 2. $l_H^* < l_L^*$.

Consider a new allocation in which we increase θ_L and θ_M by setting $w_H = w_H^* - \epsilon$ and $\tilde{w}_L = w_L^* + \epsilon p_H / p_L$, and leave the rest of the efficient allocation unchanged. This increases the total output and relaxes the incentive constraint from type H to type M and type M to type L, creating a contradiction.

Case 3. $l_L^* \leq l_H^* < l_M^*$.

Consider a new allocation in which we set $w_L = 0$, keep θ_M unchanged, i.e., $\frac{w_M}{w_H} = \frac{w_M^*}{w_H^*}$, set $l_L = 0$, and keep l_M^* and l_H^* unchanged. This strictly increases total output and decreases type L's disutility to 0.

We construct the consumption allocation such that incentive constraints hold with equality and total output is used up for consumption: $u(c_H) - v(l_H^*) = u(c_M) - v(\theta_M l_M^*)$ and $u(c_M) - v(l_M^*) = u(c_L) - v(0)$.

To see that we can find such a consumption allocation, observe that if we divide all of the output between c_M and c_H such that type H's incentive constraint holds with equality, we obtain $u(c_H) - v(l_H^*) > u(0) - v(0)$, $u(c_M) - v(l_M^*) > u(0) - v(0)$. Suppose that the first strict inequality does not hold. Then, $u(c_H^*) - v(l_H^*) \leq u(0) - v(0)$, which gives a contradiction in a similar way as in case 1. Now suppose for contradiction that $u(\tilde{c}_M) - v(l_M^*) \leq u(0) - v(0)$. Then, consider a new allocation in which we give both type L and M $(0, 0)$ and type H (c_H^*, l_H^*). This is incentive compatible, feasible, and gives a strictly higher utility, which is a contradiction. Now, decrease c_M, c_H and increase c_L until the incentive constraint of type M holds with equality, keeping the incentive constraint of type H holding with equality. In this way we construct our consumption allocation. We argue that in this allocation $c_H > c_H^*$, which also implies $c_M > c_M^*$ due to incentive constraint of type H holding with equality. Suppose that $c_H \leq c_H^*$. This implies $c_M \leq c_M^*$ and $c_L > c_L^*$ (as the total output is increased in the new allocation), which implies $u(c_M) - v(l_M^*) \leq u(c_M) - v(l_M^*) < u(c_L) - v(0)$, which contradicts the type M's incentive constraint holding with equality. Thus, $c_H > c_H^*$ and $c_M > c_M^*$.

Hence, in the perturbed allocation, type M and type H welfare is strictly increased, and type L welfare is increased because $u(c_M) - v(l_M^*) = u(c_L) - v(0) > u(c_M^*) - v(l_M^*) > u(c_L^*) - v(l_L^*)$. Thus, perturbed allocation offers an improvement over the efficient allocation, yielding the desired contradiction.

□

5 Conclusion

This paper has studied the socially optimal distribution of skills in a Mirrleesian economy. We have shown that optimal skill distribution is either perfectly equal or perfectly unequal, but an interior level of skill inequality is never optimal. We have also provided conditions on the parameters under which perfectly equal and perfectly unequal skill distributions are optimal.

We acknowledge that our analysis is purely theoretical, and we do not take a stance on any particular interpretation of the skill distribution choice. However, we believe that our model of skill distribution choice could serve as a benchmark for analyzing policy questions regarding education and SBTC. A normative analysis of that kind would require a modification and enrichment of the current model tailored to the specific policy question at hand. We believe that such an analysis of education and SBTC policies based on our model could be an interesting line of future research.

References

Boadway, R. W., and P. Pestieau (2006): "Tagging and Redistributive Taxation," *Annales d'Économie et de Statistique*, (83-84), 123–147.

Bovenberg, A. L., and B. Jacobs (2006): "Redistribtion and Education Subsidies are Siamese Twins," *Journal of Public Economics*, (85), 2005–2035.

Brett, C., and J. A. Weymark (2008): "Public Good Provision and the Comparative Statics of Optimal Nonlinear Income Taxation," *International Economic Review*, 49(1), 255–290.

Cremer, H., P. Pestieau, and M. Racionero (2010): "Unequal Wages for Equal Utilities," *International Tax and Public Finance*, forthcoming.

Hamilton, J., and P. Pestieau (2005): "Optimal Income Taxation and the Ability Distribution: Implications for Migration Equilibria," *International Tax and Public Finance*, 12(1), 29–45.

Hare, P., and D. Ulph (1979): "On Education and Distribution," *The Journal of Political Economy*, 87(5), 193–212.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

Naito, H. (1999): "Re-Examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implication for Productive Eficiency," *Journal of Public Economics*, 71, 165–188.

Simula, L. (2007): "Optimality Conditions and Comparative Static Properties of Non-linear Income Taxes Revisited," *Paris School of Economics Working Paper*, (2007-15).

Stiglitz, J. E. (1982): "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17, 213–240.

Violante, G. L. (2009): "Skill-Biased Technical Change," *The New Palgrave Dictionary of Economics*.